5 Mutual Information and Channel Capacity

In Section 2, we have seen the use of a quantity called entropy to measure the amount of randomness in a random variable. In this section, we introduce several more information-theoretic quantities. These quantities are important in the study of Shannon's results.

5.1 Information-Theoretic Quantities

Definition 5.1. Recall that, the **entropy** of a discrete random variable X is defined in Definition 2.41 to be

$$H(X) = -\sum_{x \in S_X} p_X(x) \log_2 p_X(x) = -\mathbb{E}\left[\log_2 p_X(X)\right]. \tag{16}$$

Similarly, the entropy of a discrete random variable Y is given by

$$H(Y) = -\sum_{y \in S_Y} p_Y(y) \log_2 p_Y(y) = -\mathbb{E}\left[\log_2 p_Y(Y)\right]. \tag{17}$$

In our context, the X and Y are input and output of a discrete memoryless channel, respectively. In such situation, we have introduced some new notations in Section 3.1:

$S_X \equiv \mathcal{X}$ $\quad p_X(x) \equiv p(x)$ $\quad p_{Y|X}(y|x) \equiv Q(y|x)$ ⟶ Q matrix

$S_Y \equiv \mathcal{Y}$ $\quad p_Y(y) \equiv q(y)$ $\quad p_{X,Y}(x,y) \equiv p(x,y)$ ⟶ P matrix

Under such notations, (16) and (17) become

$$H(X) = -\sum_{x \in \mathcal{X}} p(x) \log_2 p(x) = -\mathbb{E}\left[\log_2 p(X)\right] \equiv H(\underline{p}) \tag{18}$$

(the pmf of X expressed using a row vector.)

and

$$H(Y) = -\sum_{y \in \mathcal{Y}} q(y) \log_2 q(y) = -\mathbb{E}\left[\log_2 q(Y)\right]. \tag{19}$$

Definition 5.2. The **joint entropy** for two random variables X and Y is given by

$$H(X,Y) = -\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x,y) \log_2 p(x,y) = -\mathbb{E}\left[\log_2 p(X,Y)\right].$$

$\sum_{(x,y)}$ elements inside the P matrix

Example 5.3. Random variables X and Y have the following joint pmf matrix $\mathbf{P}$:

$$\mathbf{P} = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ \sum\downarrow \\ 1/4 \end{array} \begin{array}{c} \begin{array}{cccc} y_1 & y_2 & y_3 & y_4 \end{array} \\ \begin{bmatrix} \frac{1}{8} & \frac{1}{16} & \frac{1}{16} & \frac{1}{4} \\ \frac{1}{16} & \frac{1}{8} & \frac{1}{16} & 0 \\ \frac{1}{32} & \frac{1}{32} & \frac{1}{16} & 0 \\ \frac{1}{32} & \frac{1}{32} & \frac{1}{16} & 0 \end{bmatrix} \\ \begin{array}{cccc} \sum\downarrow & \sum\downarrow & \sum\downarrow & \sum\downarrow \\ 1/4 & 1/4 & 1/4 & 1/4 \end{array} \end{array} \begin{array}{c} P(x) \\ \sum \rightarrow 1/2 \\ \sum \rightarrow 1/4 \\ \sum \rightarrow 1/8 \\ \sum \rightarrow 1/8 \end{array}$$

Find $H(X)$, $H(Y)$ and $H(X,Y)$.

$H(X) = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{4}\log_2\frac{1}{4} - 2\times\frac{1}{8}\log_2\frac{1}{8} = \frac{1}{2} + \frac{2}{4} + \frac{6}{8} = \frac{7}{4}$ bits

$H(Y) = -4\times\frac{1}{4}\log_2\frac{1}{4} = 2 \qquad = \log_2 4 = 2$

$H(X,Y) = -2\times\frac{1}{8}\log_2\frac{1}{8} - 6\times\frac{1}{16}\log_2\frac{1}{16} - \frac{1}{4}\log_2\frac{1}{4} - 4\times\frac{1}{32}\log_2\frac{1}{32} - 3\times 0\log_2 0 = \frac{27}{8}$ bits

Definition 5.4. Conditional entropy:

(a) The (conditional) entropy of Y when we know $X = x$ is denoted by $H(Y|X=x)$ or simply $H(Y|x)$. It can be calculated from

$$H(Y|x) = -\sum_{y\in\mathcal{Y}} Q(y|x)\log_2 Q(y|x) = -\mathbb{E}\left[\log_2(Q(Y|x))|\, X = x\right].$$

P[Y=y|X=x]

updated (revised)

- Note that the above formula is what we should expect it to be. When we want to find the entropy of Y, we use (19):

$$H(Y) = -\sum_{y\in\mathcal{Y}} q(y)\log_2 q(y).$$

P[Y=y]

When we have an extra piece of information that $X = x$, we should update the probability about Y from the unconditional probability $q(y)$ to the conditional probability $Q(y|x)$.

- Note that when we consider $Q(y|x)$ with the value of x fixed and the value of y varied, we simply get the whole x-row from $\mathbf{Q}$ matrix. So, to find $H(Y|x)$, we simply find the "usual" entropy from the probability values in the row corresponding to x in the $\mathbf{Q}$ matrix.

(b) The (average) conditional entropy of Y when we know X is denoted by $H(Y|X)$. It can be calculated from

$$\begin{aligned} H(Y|X) &= \sum_{x\in\mathcal{X}} p(x) H(Y|x) \\ &= -\sum_{x\in\mathcal{X}} p(x) \sum_{y\in\mathcal{Y}} Q(y|x) \log_2 Q(y|x) \\ &= -\sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} p(x,y) \log_2 Q(y|x) \\ &= -\mathbb{E}\left[\log_2 Q(Y|X)\right] \end{aligned}$$

- Note that $Q(y|x) = \frac{p(x,y)}{p(x)}$. Therefore,

$$\begin{aligned} H(Y|X) &= -\mathbb{E}\left[\log_2 Q(Y|X)\right] = -\mathbb{E}\left[\log_2 \frac{p(X,Y)}{p(X)}\right] \\ &= \left(-\mathbb{E}\left[\log_2 p(X,Y)\right]\right) - \left(-\mathbb{E}\left[\log_2 p(X)\right]\right) \\ &= H(X,Y) - H(X) \end{aligned}$$

Example 5.5. Continue from Example 5.3. Random variables X and Y have the following joint pmf matrix $\mathbf{P}$:

$$\mathbf{P} = \begin{bmatrix} \frac{1}{8} & \frac{1}{16} & \frac{1}{16} & \frac{1}{4} \\ \frac{1}{16} & \frac{1}{8} & \frac{1}{16} & 0 \\ \frac{1}{32} & \frac{1}{32} & \frac{1}{16} & 0 \\ \frac{1}{32} & \frac{1}{32} & \frac{1}{16} & 0 \end{bmatrix}$$

Find $H(Y|X)$ and $H(X|Y)$.

$H(Y|X) = H(X,Y) - H(X) = \frac{27}{8} - \frac{7}{4} = \frac{13}{8}$

$H(X|Y) = H(X,Y) - H(Y) = \frac{27}{8} - 2 = \frac{11}{8}$

$$P = \begin{array}{c|cccc} x\backslash y & 1 & 2 & 3 & 4 \\ 1 & 1/8 & 1/16 & 1/16 & 1/4 \\ 2 & 1/16 & 1/8 & 1/16 & 0 \\ 3 & 1/32 & 1/32 & 1/16 & 0 \\ 4 & 1/32 & 1/32 & 1/16 & 0 \end{array} \quad \begin{array}{c} p(x) \\ 1/2 \\ 1/4 \\ 1/8 \\ 1/8 \end{array} \quad \begin{array}{c} \xrightarrow{\times 2} \\ \xrightarrow{\times 4} \\ \xrightarrow{\times 8} \\ \xrightarrow{\times 8} \end{array} \quad Q = \begin{array}{c|cccc} x\backslash y & 1 & 2 & 3 & 4 \\ 1 & 1/4 & 1/8 & 1/8 & 1/2 \\ 2 & 1/4 & 1/2 & 1/4 & 0 \\ 3 & 1/4 & 1/4 & 1/2 & 0 \\ 4 & 1/4 & 1/4 & 1/2 & 0 \end{array}$$

$$H(Y|1) = -\tfrac{1}{4}\log_2\tfrac{1}{4} - \tfrac{2}{8}\log_2\tfrac{1}{8} - \tfrac{1}{2}\log_2\tfrac{1}{2} = \tfrac{7}{4}$$
$$H(Y|2) = -\tfrac{2}{4}\log_2\tfrac{1}{4} - \tfrac{1}{2}\log_2\tfrac{1}{2} = 1 + \tfrac{1}{2} = \tfrac{3}{2}$$
$$H(Y|3) = 3/2$$
$$H(Y|4) = 3/2$$

$q(y)$ $\quad \frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4}$

$\times 4$ $\quad \times 4 \quad \times 4 \quad \times 4$

$$P_{x|y} = \begin{bmatrix} 1/2 & 1/4 & 1/4 & 1 \\ 1/4 & 1/2 & 1/4 & 0 \\ 1/8 & 1/8 & 1/4 & 0 \\ 1/8 & 1/8 & 1/4 & 0 \end{bmatrix}$$

$$Q(y|x) = P[Y=y|X=x] = \frac{P[Y=y, X=x]}{P[X=x]}$$

$$P[X=x|Y=y] = \frac{P[Y=y, X=x]}{P[Y=y]}$$

$$H(Y|X) = \tfrac{1}{2}\times\tfrac{7}{4} + \tfrac{1}{4}\times\tfrac{3}{2} + \tfrac{1}{8}\times\tfrac{3}{2} + \tfrac{1}{8}\times\tfrac{3}{2} = \tfrac{13}{8}$$

$H(X|1) = H(X|2)$ $\quad H(X|3) = 2$

$=$ $\quad H(X|4) = 0$

$7/4$

$$H(X|Y) = \tfrac{1}{4}\times\tfrac{7}{4} + \tfrac{1}{4}\times\tfrac{7}{4} + \tfrac{1}{4}\times 2 + \tfrac{1}{4}\times 0 = \tfrac{11}{8}$$

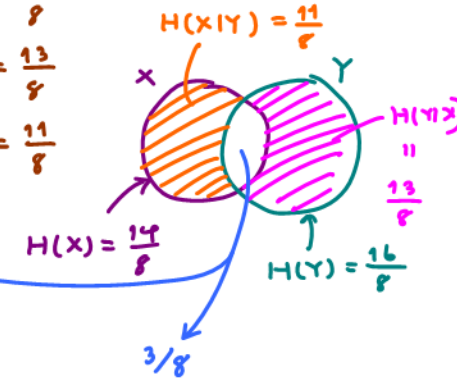
Compare these with

$$P(A\cap B) = P(A) - P(A\backslash B) = P(B) - P(B\backslash A)$$
$$= P(A) + P(B) - P(A\cup B)$$

$$I(X;Y) \equiv \begin{cases} \text{The area in the middle} \\ = H(X) - H(X|Y) \\ = H(Y) - H(Y|X) \\ = H(X) + H(Y) - H(X,Y) \end{cases}$$

$$H(X) = \tfrac{7}{4} = \tfrac{14}{8}$$
$$H(Y) = 2 = \tfrac{16}{8}$$
$$H(X,Y) = \tfrac{27}{8} = \tfrac{27}{8}$$
$$H(Y|X) = \tfrac{13}{8} = \tfrac{13}{8}$$
$$H(X|Y) = \tfrac{11}{8} = \tfrac{11}{8}$$
$$I(X;Y) = \tfrac{3}{8}$$

$$H(X,Y) = H(X) + H(Y|X)$$
$$= H(Y) + H(X|Y)$$



Definition 5.6. The **mutual information**[14] $I(X;Y)$ between two random variables X and Y is defined as

$$I(X;Y) = H(X) - H(X|Y) \tag{20}$$
$$= H(Y) - H(Y|X) \tag{21}$$
$$= H(X) + H(Y) - H(X,Y) \tag{22}$$
$$= \mathbb{E}\left[\log_2 \frac{p(X,Y)}{p(X)\,q(Y)}\right] = \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\log\frac{p(x,y)}{p(x)\,q(y)} \tag{23}$$
$$= \mathbb{E}\left[\log_2 \frac{P_{X|Y}(X|Y)}{p(X)}\right] = \mathbb{E}\left[\log_2 \frac{Q(Y|X)}{q(Y)}\right]. \tag{24}$$

- Mutual information quantifies the reduction in the uncertainty of one random variable due to the knowledge of the other.

$I(X;Y) = H(X) - H(X|Y)$ = Amount of reduction in the randomness of X due to the knowledge of Y.
↳ Amount of randomness still remained in X when we know Y.
Amount of randomness in X

- Mutual information is a measure of the amount of information one random variable contains about another [3, p 13].
- It is natural to think of $I(X;Y)$ as a measure of how far (distance) X and Y are from being independent. So, $I(X;Y)$ measures the amount of dependency btw the two RVs.
  - Technically, it is the (Kullback-Leibler) divergence between the joint and product-of-marginal distributions.

5.7. Some important properties

(a) $H(X,Y) = H(Y,X)$ and $I(X;Y) = I(Y;X)$. by the symmetry in their defn.
However, in general, $H(X|Y) \neq H(Y|X)$. by the asymmetry in their defn.

(b) I and H are always ≥ 0.

(c) There is a one-to-one correspondence between Shannon's information measures and set theory. We may use an **information diagram**, which

[14] The name mutual information and the notation $I(X;Y)$ was introduced by [Fano, 1961, Ch 2].

is a variation of a Venn diagram, to represent relationship between Shannon's information measures. This is similar to the use of the Venn diagram to represent relationship between probability measures. These diagrams are shown in Figure 7.

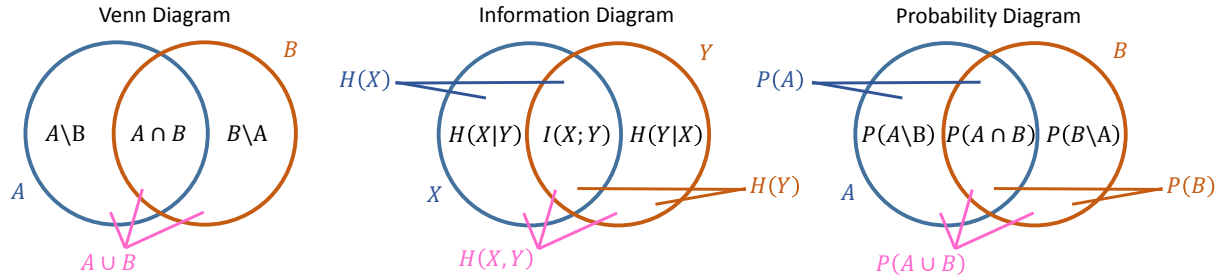


Figure 7: Venn diagram and its use to represent relationship between information measures and relationship between information measures

- Chain rule for information measures:

$$H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y).$$

(d) $I(X;Y) \geq 0$ with equality if and only if X and Y are independent.

- When this property is applied to the information diagram (or definitions (20), (21), and (22) for $I(X,Y)$), we have

  (i) $H(X|Y) \leq H(X)$,

  (ii) $H(Y|X) \leq H(Y)$,

  (iii) $H(X,Y) \leq H(X) + H(Y)$

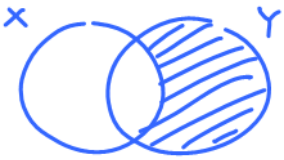
  Moreover, each of the inequalities above becomes equality if and only $X \perp\!\!\!\perp Y$.

(e) We have seen in Section 2.4 that

$$\underset{\text{deterministic (degenerated)}}{0} \leq H(X) \leq \underset{\text{uniform}}{\log_2 |\mathcal{X}|}. \tag{25}$$

Similarly,

$$\underset{\text{deterministic (degenerated)}}{0} \leq H(Y) \leq \underset{\text{uniform}}{\log_2 |\mathcal{Y}|}. \tag{26}$$



For conditional entropy, we have

$$\underset{\exists g\, Y=g(X)}{0} \leq H\left(Y\,|X\right) \leq \underset{X \perp\!\!\!\perp Y}{H\left(Y\right)} \tag{27}$$

and

$$\underset{\exists g\, X=g(Y)}{0} \leq H\left(X\,|Y\right) \leq \underset{X \perp\!\!\!\perp Y}{H\left(X\right)}. \tag{28}$$

For mutual information, we have

Think about this →

$$\underset{X \perp\!\!\!\perp Y}{0} \leq I\left(X;Y\right) \leq \underset{\exists g\, X=g(Y)}{H\left(X\right)} \tag{29}$$

and

$$\underset{X \perp\!\!\!\perp Y}{0} \leq I\left(X;Y\right) \leq \underset{\exists g\, Y=g(X)}{H\left(Y\right)}. \tag{30}$$

Combining 25, 26, 29, and 30, we have

$$0 \leq I\left(X;Y\right) \leq \min\left\{H\left(X\right), H\left(Y\right)\right\} \leq \min\left\{\log_2\left|\mathcal{X}\right|, \log_2\left|\mathcal{Y}\right|\right\} \tag{31}$$

(f) $H\left(X\,|X\right) = 0$ and $I(X;X) = H(X)$.

Example 5.8. Find the mutual information $I(X;Y)$ between the two random variables X and Y whose joint pmf matrix is given by $\mathbf{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & 0 \end{bmatrix}$.

$I(X;Y) = H(X) + H(Y) - H(X,Y) = 0.1226$

$H(X,Y) = -\frac{1}{2}\log_2\frac{1}{2} - \frac{2}{4}\log_2\frac{1}{4} = 1.5$

$H(Y) = 0.8113$

$H(X) = -\frac{3}{4}\log_2\frac{3}{4} - \frac{1}{4}\log_2\frac{1}{4} = 0.8113$

$\begin{bmatrix} 1/2 & 1/4 \\ 1/4 & 0 \end{bmatrix} \to \begin{matrix} 3/4 \\ 1/4 \end{matrix}$

$\downarrow \quad \downarrow$

$3/4 \quad 1/4$

Example 5.9. Find the mutual information $I(X;Y)$ between the two random variables X and Y whose $\underline{\mathbf{p}} = \left[\frac{1}{4}, \frac{3}{4}\right]$ and $\mathbf{Q} = \begin{bmatrix} \frac{1}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{1}{4} \end{bmatrix}$.

$\times 1/4 \to \begin{bmatrix} \frac{1}{16} & \frac{3}{16} \\ \frac{9}{16} & \frac{3}{16} \end{bmatrix} = P$

$\times 3/4 \to$

$I(X;Y) = 0.1432$

$P = \begin{bmatrix} \frac{1}{4} & \frac{3}{4} \end{bmatrix}$ $\qquad$ $Q = \begin{array}{c|cc} x \backslash y & 0 & 1 \\ 0 & \frac{1}{4} & \frac{3}{4} \\ 1 & \frac{3}{4} & \frac{1}{4} \end{array}$

Method 1 : $I(X;Y) = H(X) + H(Y) - H(X,Y)$

First, convert the given information into the joint pmf matrix.

$P = \begin{bmatrix} \frac{1}{4} & \frac{3}{4} \end{bmatrix} \longrightarrow H(X) = H\left(\begin{bmatrix} \frac{1}{4} & \frac{3}{4} \end{bmatrix}\right) = 0.8113$

$Q = \begin{array}{c|cc} x \backslash y & 0 & 1 \\ 0 & \frac{1}{4} & \frac{3}{4} \\ 1 & \frac{3}{4} & \frac{1}{4} \end{array}$ $\xrightarrow[\times \frac{3}{4}]{\times \frac{1}{4}}$ $\begin{array}{c|cc} x \backslash y & 0 & 1 \\ 0 & \frac{1}{16} & \frac{3}{16} \\ 1 & \frac{9}{16} & \frac{3}{16} \end{array} = P$ $\longrightarrow H(X,Y) = H\left(\left[\frac{1}{16}, \frac{3}{16}, \frac{9}{16}, \frac{3}{16}\right]\right) = 1.6226$

(sum along each column)

$\frac{10}{16}$ $\quad$ $\frac{6}{16}$ $\longrightarrow H(Y) = H\left(\left[\frac{5}{8}, \frac{3}{8}\right]\right) = 0.9544$

$I(X;Y) = H(X) + H(Y) - H(X,Y) \approx 0.1432.$

Method 2 : $I(X;Y) = H(Y) - H(Y|X)$

$H(Y|X) = \sum_x p(x) \underbrace{H(Y|x)}_{0.8113} = 0.8113 \times \overbrace{\sum_x p(x)}^{1} = 0.8113$

0.8113 ← Each row of Q contains $\frac{1}{4}, \frac{3}{4}$. So, $H(Y|x) = 0.8113$ for any x. (for any row.)

Same as Ex. 5.8

$I(X;Y) = \underbrace{H(Y)}_{0.9544} - H(Y|X) = 0.1432$

To find $H(Y)$, we need $q(y)$

$q = pQ = \begin{bmatrix} \frac{1}{4} & \frac{3}{4} \end{bmatrix} \begin{bmatrix} \frac{1}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{1}{4} \end{bmatrix}$

$= \begin{bmatrix} \frac{10}{16} & \frac{6}{16} \end{bmatrix} = \begin{bmatrix} \frac{5}{8} & \frac{3}{8} \end{bmatrix}$